

#5
LMS
3-15-01

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,198

DATE: 01/04/2001
TIME: 11:31:43

Input Set : A:\Pto.amc
Output Set: N:\CRF3\01042001\I673198.raw

2 <110> APPLICANT: MIYAKE, Koichiro; HASHIMOTO, Shinichi; MOTOYAMA Hiroaki;
3 OZAKI, Akio; SETO, Haruo; KUZAYAMA, Tomohisa; TAKAHASHI, Shunji
5 <120> TITLE OF INVENTION: A process for producing isoprenoid compounds by
6 microorganisms and a method for screening compounds with
7 antibiotic or weeding activity
W--> 9 <130> FILE REFERENCE:
C--> 11 <140> CURRENT APPLICATION NUMBER: US/09/673,198
C--> 12 <141> CURRENT FILING DATE: 2000-10-12
14 <150> PRIOR APPLICATION NUMBER: JP98/103101
15 <151> PRIOR FILING DATE: 1998-04-14
17 <150> PRIOR APPLICATION NUMBER: JP98/221910
18 <151> PRIOR FILING DATE: 1998-08-05
20 <150> PRIOR APPLICATION NUMBER: JP99/035739
21 <151> PRIOR FILING DATE: 1999-02-15
23 <160> NUMBER OF SEQ ID NOS: 34
25 <170> SOFTWARE: PatentIn Ver. 2.0
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 620
29 <212> TYPE: PRT
30 <213> ORGANISM: Escherichia coli
32 <400> SEQUENCE: 1
33 Met Ser Phe Asp Ile Ala Lys Tyr Pro Thr Leu Ala Leu Val Asp Ser
34 1 5 10 15
36 Thr Gln Glu Leu Arg Leu Leu Pro Lys Glu Ser Leu Pro Lys Leu Cys
37 20 25 30
39 Asp Glu Leu Arg Arg Tyr Leu Leu Asp Ser Val Ser Arg Ser Ser Gly
40 35 40 45
42 His Phe Ala Ser Gly Leu Gly Thr Val Glu Leu Thr Val Ala Leu His
43 50 55 60
45 Tyr Val Tyr Asn Thr Pro Phe Asp Gln Leu Ile Trp Asp Val Gly His
46 65 70 75 80
48 Gln Ala Tyr Pro His Ile Leu Thr Gly Arg Arg Asp Lys Ile Gly
49 85 90 95
51 Thr Ile Arg Gln Lys Gly Gly Leu His Pro Phe Pro Trp Arg Gly Glu
52 100 105 110
54 Ser Glu Tyr Asp Val Leu Ser Val Gly His Ser Ser Thr Ser Ile Ser
55 115 120 125
57 Ala Gly Ile Gly Ile Ala Val Ala Ala Glu Lys Glu Gly Lys Asn Arg
58 130 135 140
60 Arg Thr Val Cys Val Ile Gly Asp Gly Ala Ile Thr Ala Gly Met Ala
61 145 150 155 160
63 Phe Glu Ala Met Asn His Ala Gly Asp Ile Arg Pro Asp Met Leu Val
64 165 170 175
66 Ile Leu Asn Asp Asn Glu Met Ser Ile Ser Glu Asn Val Gly Ala Leu
67 180 185 190
69 Asn Asn His Leu Ala Gln Leu Leu Ser Gly Lys Leu Tyr Ser Ser Leu
70 195 200 205

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,198

DATE: 01/04/2001
TIME: 11:31:43

Input Set : A:\Pto.amc
Output Set: N:\CRF3\01042001\I673198.raw

72 Arg Glu Gly Lys Lys Val Phe Ser Gly Val Pro Pro Ile Lys Glu
73 210 215 220
75 Leu Leu Lys Arg Thr Glu Glu His Ile Lys Gly Met Val Val Pro Gly
76 225 230 235 240
78 Thr Leu Phe Glu Glu Leu Gly Phe Asn Tyr Ile Gly Pro Val Asp Gly
79 245 250 255
81 His Asp Val Leu Gly Leu Ile Thr Thr Leu Lys Asn Met Arg Asp Leu
82 260 265 270
84 Lys Gly Pro Gln Phe Leu His Ile Met Thr Lys Lys Gly Arg Gly Tyr
85 275 280 285
87 Glu Pro Ala Glu Ile Asp Pro Ile Thr Phe His Ala Val Pro Lys Phe
88 290 295 300
90 Asp Pro Ser Ser Gly Cys Leu Pro Lys Ser Ser Gly Gly Leu Pro Ser
91 305 310 315 320
93 Tyr Ser Lys Ile Phe Gly Asp Trp Leu Cys Glu Thr Ala Ala Lys Asp
94 325 330 335
96 Asn Lys Leu Met Ala Ile Thr Pro Ala Met Arg Glu Gly Ser Gly Met
97 340 345 350
99 Val Glu Phe Ser Arg Lys Phe Pro Asp Arg Tyr Phe Asp Val Ala Ile
100 355 360 365
102 Ala Glu Gln His Ala Val Thr Phe Ala Ala Gly Leu Ala Ile Gly Gly
103 370 375 380
105 Tyr Lys Pro Ile Val Ala Ile Tyr Ser Thr Phe Leu Gln Arg Ala Tyr
106 385 390 395 400
108 Asp Gln Val Leu His Asp Val Ala Ile Gln Lys Leu Pro Val Leu Phe
109 405 410 415
111 Ala Ile Asp Arg Ala Gly Ile Val Gly Ala Asp Gly Gln Thr His Gln
112 420 425 430
114 Gly Ala Phe Asp Leu Ser Tyr Leu Arg Cys Ile Pro Glu Met Val Ile
115 435 440 445
117 Met Thr Pro Ser Asp Glu Asn Glu Cys Arg Gln Met Leu Tyr Thr Gly
118 450 455 460
120 Tyr His Tyr Asn Asp Gly Pro Ser Ala Val Arg Tyr Pro Arg Gly Asn
121 465 470 475 480
123 Ala Val Gly Val Glu Leu Thr Pro Leu Glu Lys Leu Pro Ile Gly Lys
124 485 490 495
126 Gly Ile Val Lys Arg Arg Gly Glu Lys Leu Ala Ile Leu Asn Phe Gly
127 500 505 510
129 Thr Leu Met Pro Glu Ala Ala Lys Val Ala Glu Ser Leu Asn Ala Thr
130 515 520 525
132 Leu Val Asp Met Arg Phe Val Lys Pro Leu Asp Glu Ala Leu Ile Leu
133 530 535 540
135 Glu Met Ala Ala Ser His Glu Ala Leu Val Thr Val Glu Glu Asn Ala
136 545 550 555 560
138 Ile Met Gly Gly Ala Gly Ser Gly Val Asn Glu Val Leu Met Ala His
139 565 570 575
141 Arg Lys Pro Val Pro Val Leu Asn Ile Gly Leu Pro Asp Phe Phe Ile
142 580 585 590
144 Pro Gln Gly Thr Gln Glu Glu Met Arg Ala Glu Leu Gly Leu Asp Ala

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,198

DATE: 01/04/2001
TIME: 11:31:43

Input Set : A:\Pto.amc
Output Set: N:\CRF3\01042001\I673198.raw

145 595 600 605
147 Ala Gly Met Glu Ala Ala Lys Ile Lys Ala Trp Leu Ala
148 610 615 620
151 <210> SEQ ID NO: 2
152 <211> LENGTH: 299
153 <212> TYPE: PRT
154 <213> ORGANISM: Escherichia coli
155 <400> SEQUENCE: 2
156 Met Asp Phe Pro Gln Gln Leu Glu Ala Cys Val Lys Gln Ala Asn Gln
157 1 5 10 15
158 Ala Leu Ser Arg Phe Ile Ala Pro Leu Pro Phe Gln Asn Thr Pro Val
159 20 25 30
160 Val Glu Thr Met Gln Tyr Gly Ala Leu Leu Gly Gly Lys Arg Leu Arg
161 35 40 45
162 Pro Phe Leu Val Tyr Ala Thr His Met Phe Gly Val Ser Thr Asn
163 50 55 60
164 Thr Leu Asp Ala Pro Ala Ala Val Glu Cys Ile His Ala Tyr Ser
165 65 70 75 80
166 Leu Ile His Asp Asp Leu Pro Ala Met Asp Asp Asp Asp Leu Arg Arg
167 85 90 95
168 Gly Leu Pro Thr Cys His Val Lys Phe Gly Glu Ala Asn Ala Ile Leu
169 100 105 110
170 Ala Gly Asp Ala Leu Gln Thr Leu Ala Phe Ser Ile Leu Ser Asp Ala
171 115 120 125
172 Asp Met Pro Glu Val Ser Asp Arg Asp Arg Ile Ser Met Ile Ser Glu
173 130 135 140
174 Leu Ala Ser Ala Ser Gly Ile Ala Gly Met Cys Gly Gly Gln Ala Leu
175 145 150 155 160
176 Asp Leu Asp Ala Glu Gly Lys His Val Pro Leu Asp Ala Leu Glu Arg
177 165 170 175
178 Ile His Arg His Lys Thr Gly Ala Leu Ile Arg Ala Ala Val Arg Leu
179 180 185 190
180 Gly Ala Leu Ser Ala Gly Asp Lys Gly Arg Arg Ala Leu Pro Val Leu
181 195 200 205
182 Asp Lys Tyr Ala Glu Ser Ile Gly Leu Ala Phe Gln Val Gln Asp Asp
183 210 215 220
184 Ile Leu Asp Val Val Gly Asp Thr Ala Thr Leu Gly Lys Arg Gln Gly
185 225 230 235 240
186 Ala Asp Gln Gln Leu Gly Lys Ser Thr Tyr Pro Ala Leu Leu Gly Leu
187 245 250 255
188 Glu Gln Ala Arg Lys Lys Ala Arg Asp Leu Ile Asp Asp Ala Arg Gln
189 260 265 270
190 Ser Leu Lys Gln Leu Ala Glu Gln Ser Leu Asp Thr Ser Ala Leu Glu
191 275 280 285
192 Ala Leu Ala Asp Tyr Ile Ile Gin Arg Asn Lys
193 290 295
194 <210> SEQ ID NO: 3
195 <211> LENGTH: 80
196 <212> TYPE: PRT

RAW SEQUENCE LISTING DATE: 01/04/2001
 PATENT APPLICATION: US/09/673,198 TIME: 11:31:43

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\01042001\I673198.raw

218 <213> ORGANISM: Escherichia coli
 220 <400> SEQUENCE: 3
 221 Met Pro Lys Lys Asn Glu Ala Pro Ala Ser Phe Glu Lys Ala Leu Ser
 222 1 5 10 15
 224 Glu Leu Glu Gln Ile Val Thr Arg Leu Glu Ser Gly Asp Leu Pro Leu
 225 20 25 30
 227 Glu Glu Ala Leu Asn Glu Phe Glu Arg Gly Val Gln Leu Ala Arg Gln
 228 35 40 45
 230 Gly Gln Ala Lys Leu Gln Gln Ala Glu Gln Arg Val Gln Ile Leu Leu
 231 50 55 60
 233 Ser Asp Asn Glu Asp Ala Ser Leu Thr Pro Phe Thr Pro Asp Asn Glu
 234 65 70 75 80
 237 <210> SEQ ID NO: 4
 238 <211> LENGTH: 348
 239 <212> TYPE: PRT
 240 <213> ORGANISM: Escherichia coli
 242 <400> SEQUENCE: 4
 243 Val Thr Gly Val Asn Glu Cys Ser Arg Ser Thr Cys Asn Leu Lys Tyr
 244 1 5 10 15
 246 Asp Glu Tyr Ser Arg Ser Gly Ser Met Gln Tyr Asn Pro Leu Gly Lys
 247 20 25 30
 249 Thr Asp Leu Arg Val Ser Arg Leu Cys Leu Gly Cys Met Thr Phe Gly
 250 35 40 45
 252 Glu Pro Asp Arg Gly Asn His Ala Trp Thr Leu Pro Glu Glu Ser Ser
 253 50 55 60
 255 Arg Pro Ile Ile Lys Arg Ala Leu Glu Gly Ile Asn Phe Phe Asp
 256 65 70 75 80
 258 Thr Ala Asn Ser Tyr Ser Asp Gly Ser Ser Glu Glu Ile Val Gly Arg
 259 85 90 95
 261 Ala Leu Arg Asp Phe Ala Arg Arg Glu Asp Val Val Ala Thr Lys
 262 100 105 110
 264 Val Phe His Arg Val Gly Asp Leu Pro Glu Gly Leu Ser Arg Ala Gln
 265 115 120 125
 267 Ile Leu Arg Ser Ile Asp Asp Ser Leu Arg Arg Leu Gly Met Asp Tyr
 268 130 135 140
 270 Val Asp Ile Leu Gln Ile His Arg Trp Asp Tyr Asn Thr Pro Ile Glu
 271 145 150 155 160
 273 Glu Thr Leu Glu Ala Leu Asn Asp Val Val Lys Ala Gly Lys Ala Arg
 274 165 170 175
 276 Tyr Ile Gly Ala Ser Ser Met His Ala Ser Gln Phe Ala Gln Ala Leu
 277 180 185 190
 279 Glu Leu Gln Lys Gln His Gly Trp Ala Gln Phe Val Ser Met Gln Asp
 280 195 200 205
 282 His Tyr Asn Leu Ile Tyr Arg Glu Glu Glu Arg Glu Met Leu Pro Leu
 283 210 215 220
 285 Cys Tyr Gln Glu Gly Val Ala Val Ile Pro Trp Ser Pro Leu Ala Arg
 286 225 230 235 240
 288 Gly Arg Leu Thr Arg Pro Trp Gly Glu Thr Thr Ala Arg Leu Val Ser
 289 245 250 255

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,198

DATE: 01/04/2001
TIME: 11:31:43

Input Set : A:\Pto.amc
Output Set: N:\CRF3\01042001\1673198.raw

291 Asp Glu Val Gly Lys Asn Leu Tyr Lys Glu Ser Asp Glu Asn Asp Ala
292 260 265 270
294 Gln Ile Ala Glu Arg Leu Thr Gly Val Ser Glu Glu Leu Gly Ala Thr
295 275 280 285
297 Arg Ala Gln Val Ala Leu Ala Trp Leu Leu Ser Lys Pro Gly Ile Ala
298 290 295 300
300 Ala Pro Ile Ile Gly Thr Ser Arg Glu Glu Glu Leu Asp Glu Leu Leu
301 305 310 315 320
303 Asn Ala Val Asp Ile Thr Leu Lys Pro Glu Gln Ile Ala Glu Leu Glu
304 325 330 335
306 Thr Pro Tyr Lys Pro His Pro Val Val Gly Phe Lys
307 340 345
310 <210> SEQ ID NO: 5
311 <211> LENGTH: 398
312 <212> TYPE: PRT
313 <213> ORGANISM: Escherichia coli.
315 <400> SEQUENCE: 5
316 Met Lys Gln Leu Thr Ile Leu Gly Ser Thr Gly Ser Ile Gly Cys Ser
317 1 5 10 15
319 Thr Leu Asp Val Val Arg His Asn Pro Glu His Phe Arg Val Val Ala
320 20 25 30
322 Leu Val Ala Gly Lys Asn Val Thr Arg Met Val Glu Gln Cys Leu Glu
323 35 40 45
325 Phe Ser Pro Arg Tyr Ala Val Met Asp Asp Glu Ala Ser Ala Lys Leu
326 50 55 60
328 Leu Lys Thr Met Leu Gln Gln Gly Ser Arg Thr Glu Val Leu Ser
329 65 70 75 80
331 Gly Gln Gln Ala Ala Cys Asp Met Ala Ala Leu Glu Asp Val Asp Gln
332 85 90 95
334 Val Met Ala Ala Ile Val Gly Ala Ala Gly Leu Leu Pro Thr Leu Ala
335 100 105 110
337 Ala Ile Arg Ala Gly Lys Thr Ile Leu Leu Ala Asn Lys Glu Ser Leu
338 115 120 125
340 Val Thr Cys Gly Arg Leu Phe Met Asp Ala Val Lys Gln Ser Lys Ala
341 130 135 140
343 Gln Leu Leu Pro Val Asp Ser Glu His Asn Ala Ile Phe Gln Ser Leu
344 145 150 155 160
346 Pro Gln Pro Ile Gln His Asn Leu Gly Tyr Ala Asp Leu Glu Gln Asn
347 165 170 175
349 Gly Val Val Ser Ile Leu Leu Thr Gly Ser Gly Gly Pro Phe Arg Glu
350 180 185 190
352 Thr Pro Leu Arg Asp Leu Ala Thr Met Thr Pro Asp Gln Ala Cys Arg
353 195 200 205
355 His Pro Asn Trp Ser Met Gly Arg Lys Ile Ser Val Asp Ser Ala Thr
356 210 215 220
358 Met Met Asn Lys Gly Leu Glu Tyr Ile Glu Ala Arg Trp Leu Phe Asn
359 225 230 235 240
361 Ala Ser Ala Ser Gln Met Glu Val Leu Ile His Pro Gln Ser Val Ile
362 245 250 255

VERIFICATION SUMMARY DATE: 01/04/2001
PATENT APPLICATION: US/09/673,198 TIME: 11:31:44

Input Set :: A:\Pto.amc
Output Set: N:\CRF3\01042001\I673198.raw

L:9 M:201 W: Mandatory field data missing, FILE REFERENCE
L:11 M:270 C: Current Application Number differs, Replaced Current Application Number
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1054 M:361 W: Invalid Split Codon, Sequence data for SEQ ID# 11
L:1055 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1060 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1065 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1070 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1075 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1080 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1085 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1090 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1095 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1100 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1105 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1110 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1115 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1120 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1125 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1130 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1135 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1140 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1145 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1150 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1155 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1160 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1165 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1170 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1175 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1179 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1184 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1189 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1194 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1200 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1205 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1210 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1215 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1220 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1225 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1230 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1235 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1240 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1245 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1250 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1255 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1260 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1265 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1270 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11

VERIFICATION SUMMARY DATE: 01/04/2001
PATENT APPLICATION: US/09/673,198 TIME: 11:31:44

Input Set : A:\Pto.amc
Output Set: N:\CRF3\01042001\I673198.raw

L:1275 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1280 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1285 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1290 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1295 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:1300 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11

Serial Number: 1673,198**ENTERED**

Changed a file from non-ASCII to ASCII

Changed the margins in cases where the sequence text was "wrapped" down to the next line.

Edited a format error in the Current Application Data section, specifically:

Edited the Current Application Data section with the actual current number. The number inputted by the applicant was the prior application data; or other _____.

Added the mandatory heading and subheadings for "Current Application Data".

Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.

Changed the spelling of a mandatory field (the headings or subheadings), specifically:

Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were:

Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.

Inserted colons after headings/subheadings. Headings edited included:

Deleted extra, invalid, headings used by an applicant, specifically:

Deleted: non-ASCII "garbage" at the beginning/end of files; secretary initials/filename at end of file
 page numbers throughout text; other invalid text, such as _____

Inserted mandatory headings, specifically:

Corrected an obvious error in the response, specifically:

Edited identifiers where upper case is used but lower case is required, or vice versa.

Corrected an error in the Number of Sequences field, specifically:

A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.

Deleted *ending* stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (erroneous due to a PatentIn bug). Sequences corrected:

Other:

Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.